Group members-Sri Ram, Mithilesh Reddy Maddi

Name of Project: FIFA World cup Analysis

Group Number-12

Description- The World Cup is a global football competition between the national teams of the members of FIFA, which is the sport's international governing body. Since we are all aware that FIFA is one of the most popular sports in the world, we will be looking at players' and matches' data to analyze and make a prediction.

Data Source – Data set is taken from github, data is about FIFA world cup football

DataLink- https://github.com/davidcamilo0710/QATAR\_2022\_Prediction/tree/master/data

|  |  |
| --- | --- |
| Data tables | Rows and columns |
| Qatar\_group\_stage.csv | 48, 3 |
| Squad\_stats.csv | 32, 3 |
| TeamScore.csv | 32, 7 |
| Team 1 training.csv | 9321, 7 |
| Team 2 training.csv | 9321, 7 |
| Fifa21.csv | 18944, 8 |
| Int\_matches away.csv | 23921, 15 |
| Int\_matches home.csv | 23921, 15 |
| Fifa22\_22.csv | 15847, 20 |

Fifa22\_22 Description

|  |  |
| --- | --- |
| Column Name | Description |
| Date | Date of the match played |
| home\_team | Name of the home team |
| away\_team | Name of away team |
| home\_team\_continent,  away\_team\_continent | Column D and E contains names of continent the team belong to |
| home\_team\_fifa\_rank  away\_team\_fifa\_rank | Rank of the teams |
| home\_team\_total\_fifa\_points  away\_team\_total\_fifa\_points | Points in the FIFA contest to the teams |
| home\_team\_score  away\_team\_score | Score of teams |
| tournament | Name of the tournament they played |
| City, country,  neutral\_location | Location of the match |
| home\_team\_result | Result of the match |

KEY PERFOMANCE INDICATORS

1. TEAM 1 VERSE TEAM 2 FIFA RANKING ABILITIES AND STRENGTH. (T1 VS T2) FOR THE SQUAD TEAMS.

DIAGRAM 1

At 20,805, Qatar had the highest Sum of Team1\_FIFA\_RANK and was 2,481.27% higher than Brazil, which had the lowest Sum of Team1\_FIFA\_RANK at 806.﻿ Sum of Team1\_FIFA\_RANK and total Sum of Team2\_FIFA\_RANK are positively correlated with each other.﻿ ﻿Qatar accounted for 11.24% of Sum of Team1\_FIFA\_RANK.﻿ ﻿Sum of Team1\_FIFA\_RANK and Sum of Team2\_FIFA\_RANK diverged the most for was Qatar. The total Sum of Team1\_FIFA\_RANK was 8195 higher than Sum of Team2\_FIFA\_RANK.

DIAGRAM 2

At 24987, Mexico had the highest Sum of Team1\_Offense and was 325.82% higher than Serbia, which had the lowest Sum of Team1\_Offense at 5868. ﻿Mexico accounted for 5.30% of Sum of Team1\_Offense.﻿ Sum of Team1\_Offense and Sum of Team2\_Offense diverged the most for USA. The Sum of Team1\_Offense was 12,475 higher than Sum of Team2\_Offense

1. OVERAL ABILITY & POTENTIAL OF THE SQUAD TEAMS

At 85.64, Brazil had the highest Sum of overall ability and was 20.93% higher than IR Iran, which had the lowest Sum of overall ability at 70.82. ﻿Sum of overall and total Sum of potential are positively correlated with each other. ﻿Brazil accounted for 3.40% of Sum of overall. ﻿Sum of potential and Sum of overall diverged the most when the most for USA, The combined Sum of potential is 7.46 higher than Sum of overall ability

1. SQUAD NATIONALITY COMPOSITION IN VALUE AND TOTAL EARNINGS PLUS REALEASE CLAUSE ASSESMENT

At 10,739,223,000, Spain had the highest Sum of release\_clause\_eur and was 15,106.63% higher than Costa Rica, which had the lowest Sum of release\_clause\_eur at 70,622,000. ﻿Sum of release\_clause\_eur and total Sum of value\_eur are positively correlated with each other. Spain accounted for 14.99% of Sum of release\_clause\_eur. ﻿Across all 31 nationality\_name, Sum of release\_clause\_eur ranged from 70,622,000 to 10,739,223,000. TOTAL SUM of value\_eur ranged from 41,775,000 to 5,682,495,000 WHEREAS, THE TOTAL Sum of wage\_eur (SALARY) ranged from 130,000 to 17,407,050.

1. SQUAD PLAYERS BASIC CHARACTERISTICS, (AGE, WEIGHT, HEIGHT) DISTRIBUTION

Shown using cards

In this table first table tells us which player belongs to which country he belongs to and the players overall rating. Second table is about the players present value and how much his league team is paying him and if some other team wants him how much they have to pay to release the player. We have used a slicer in this page so we can easily check player by his home team.

1. CLUB EXPENDITURE, VALUE AND WAGES

In this page a pie chart is used to show how international teams have scored highest number of goals in international matches, and a decomposition tree is which shows squad stats average which then decomposed into teams and have average player rating in the country and number of players in the country were applied in tooltips

1. PLAYERS OVERALL RATING PER NATIONALITY FOR THE SQUAD ON A SLICER USING NATIONALITIES

In this page we have total number of players in league and the average rating of the league team player rating. Other table tells us that player wage and over all rating depend on position for all the leagues.

1. PLAYER STATS BY AGE

This page we calculated average age of players that are playing and found that players in the age of early 20 to late 30 are in their top of the game, and also calculated average of players overall by age which tells us that when players age and their performance indicator.

1. CHANGE IN WAGE AND OVER ALL BY PLAYER

In this page We have created table with shot name and change in overall from past year and wage difference from past year and the graph represent average change of wage change by overall change and we know that if the person overall decreased then the wage will reduce, and other graph tells us that the average wage with player overall

Measures and calculated columns created by Us

|  |  |  |  |
| --- | --- | --- | --- |
| No | Measure name | type | Done by |
| 1 | Age of players | Measure | Sriram |
| 2 | Average overall of player | Measure | Sriram |
| 3 | Average player rating | Measure | Sriram |
| 4 | AverageWage | Measure | Sriram |
| 5 | Players In Leagues | Measure | Sriram |
| 6 | Squad\_stats\_Average | Measure | Mithilesh |
| 7 | Top20players overall | Measure | Mithilesh |
| 8 | Total goals Team Scored | Measure | Mithilesh |
| 9 | TotalTeamScores | Measure | Mithilesh |
| 10 | ChangeInOverall | Column | Sriram |
| 11 | ChangeInWage | Column | Mithilesh |

## VISUAL INSIGHTS:

Mithilesh Maddi (1,3,6,8)

As you can see in the Last T Score, I have used two distinct visuals in my presentation. The first is a table and the second is a map. In the table, the last team score is represented by all teams that participated and has offensive, midfield, and defensive scores. According to the second visual map, different countries are represented by different teams

On page 3, Using a line chart, we can see the SQUAD POTENTIAL BY NATIONALITY. At 85.64, Brazil has the highest Sum of Overall Ability, which is 20.93% higher than IR Iran at 70.82. The Sum of the overall and total Sum of potential are positively correlated. As a percentage of the overall Sum, Brazil accounted for 3.40%. USA's combined potential is 7.46 higher than its combined overall ability when the most diverged for the USA,

A two-visual representation of country-wise average can be found on page 6. There are two visuals: a pie chart divided by total team goals and the home team, and a decomposition tree which explains the average squad stats by nationality.

There are three visuals done on page 8, which is players' stats by age it has one slicer and two different columns clustered chart. Player names represent the slicer, and one clustered chart illustrates the age of players, and the other represents the average age of people.

SRIRAM (2,5,7,9)

In page 2 The two visuals on page 2 are both line charts. The first visual displays team1 and team2's FIFA ranking based on nationality. It can be seen that Qatar had the highest Sum of Team1\_FIFA\_RANK at 20,805 and was 2,481.27% higher than Brazil, which had the lowest Sum at 806. The Sum of Team1\_FIFA\_RANK and the total Sum of Team2\_FIFA\_RANK positively correlate. Team1\_FIFA\_RANK accounted for 11.24% of Qatar. The most significant divergence was for Qatar in Team1\_FIFA\_RANK and Team2\_FIFA\_RANK. As of 24987, Mexico had the highest Sum of Team1\_Offense and was 325.82% higher than Serbia, which had the lowest Sum of Team1\_Offense at 5868. The Sum of Team1\_Offense was 5.30% Mexico. Teams1\_Offense and Teams2\_Offense diverged the most for the USA. The Sum of Team1\_Offense was 12,475 higher than the Sum of Team2\_Offense.

There are three visuals on page 5 of the report. Two of them are tables, and one is a slicer. This slicer represents names of nationalities, and the tables represent the players' overall performance according to their respective countries. Messi from Argentina is the top player, and in the subsequent table, we can see details of each player's income.

In page 7, two visual representations illustrate the average player rating as per league and the second visual representation that indicates players' positions concerning their average rating and compensation. In the first visual representation, there is a line chart; however, the graph illustrates the average player rating as per league.

On page 9, we used three different types of visuals to illustrate our point. There are three types of visual representations in the first one: a table that describes players' overall change compared with their change in wage; the second one shows a stacked column chart that illustrates the change in general with change in compensation; the third is a column chart that shows the average wage in euro by an overall difference in pay.

**Group Visual Insight**

There are four visuals on page no 4 in total; three are cards, and three are line charts and clustered charts. According to the first card, the average age of the player is 25 years of age. In contrast, according to the second chart, the average weight of the player is 75.23 kilograms, and according to the third chart, the average height of the player is 181.82 kilograms.

The last visual is a line-clustered chart displaying the national name and the Sum of release clauses with a total sum value. Spain had the highest Sum of release clause\_eur at 10,739,223,000, which was 15,106.63% higher than Costa Rica, which had the lowest Sum of release clause\_eur at 70,622,000. A positive correlation exists between the Sum of release\_clause\_eur and the Sum of value\_eur. Spain accounted for a total of 14.99% of the release clauses. There was a range of 70,622,000 to 10,739,223,000 for the Sum of release\_clause\_eur across each of the 31 nationality\_names. A total of 41,775,000 to 5,682,495,000 euros were spent on value\_eur, whereas a total of 130,000 to 17,407,050 euros were spent on wage\_eur (SALARY).

**OVERALL IMPLICATIONS AND RECOMMENDATION**

FIFA World cup depends on many things like offence, midfield, defense, Goalkeeper. After analyzing the previous data on the scores of these categories we made our first graph, Argentina has the highest offense score and from the last year player positions Messi from Argentina have the highest overall for any player in the tournament and also, he has the highest league overall, from age graph we can say that many players after age 38 were mostly retiring or the performance is reduced, In second position player of Poland R. Lewandowski is in second position with 91 points overall. We predict Brazil will be the favorite to win the world cup, with Argentina, England and Spain listed as the next in line. These are the top 5 teams in the world, and Brazil has the highest chance of winning the world cup and the least likely.